

UNIVERSITI TEKNOLOGI MARA

COMPARISON OF PARAMETER ESTIMATION METHODS WHEN MULTICOLLINEARITY AND OUTLIER EXISTS

**AIDA NURASIKIN BINTI JAMIL
MUHAMMAD FAHMI BIN ABDUL MULUK
NUR SABRINA BINTI ANUAR
MOHAMAD SUFFIAN BIN ABU BAKAR**

Dissertation submitted in fulfillment
of the requirements for the degree of
Bachelor of Science (Hons) Statistics

Faculty of Computer and Mathematical Sciences

July 2019

APPROVED BY:

.....

(MRS NASUHAR BINTI AB. AZIZ)

Supervisor

Faculty of Computer and Mathematical Sciences

ABSTRACT

Ordinary Least Squares (OLS) estimator become worse in the presence of multicollinearity and outlier. Here, three methods are suggested to improve the model when multicollinearity and outlier exists, the first one is Jackknife Regression (JR) based on left out method, the second is Ridge Regression (RR) based on the addition of shrinking parameter, and the third is Latent Root Regression (LRR) by adding the latent root and latent vector. In the application, model parameters, standard errors, length of confidence intervals (L.C.I), coefficients of determination (R^2), and mean square error (MSE) of these methods are estimated. Next, the performance of these three methods are compared with OLS by using the MSE and R^2 . Based on the analysis, LRR method was the best method compared to other methods since the value of MSE is less and R^2 is higher among others. The LRR was not only the best method when multicollinearity exist, but also was the best when the presence of both multicollinearity and outlier

Keywords: Jackknife Regression (JR), Latent Root Regression (LRR), Multicollinearity, Outlier, Ridge Regression (RR)

ACKNOWLEDGMENT

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout our research work to complete the research successfully. We would like to express our deep and sincere gratitude to our research supervisor, Mrs. Nasuhar binti Ab. Aziz, who is a lecturer of Faculty Computer and Mathematical Sciences, for giving us the opportunity to do research and providing invaluable guidance throughout this research. Her dynamism, vision, sincerity and motivation have deeply inspired us. She has taught us the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under her guidance for a year. Besides, there are many people who deserve a thousand of thanks for their help and support, both spiritual and words in completing our final year project from the beginning till we finish our final year project. Furthermore, we would like to thank our beloved parents for their pray and support. We are grateful for all the motivation, encouragement and what that had been given for us for this entire time so it will be our source of strength.

In addition, we also feel blessed to have classmates who also give their help, support guidance and willing to share their knowledge through this project. Thanks to each member in our group because willing to contribute their time and energy in completing this task. Without the teamwork, we would not be able to complete the project successfully.

In conclusion, we hope that our final year project would give many benefits to others and the next generation in the future endeavours. Thank you.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER ONE: INTRODUCTION	1
1.1 Background of the Study	3
1.2 Importance and Motivation of the Study	4
1.3 Research Objectives	4
1.4 Scope and Limitation	
CHAPTER TWO: LITERATURE REVIEW	5
2.1 Multicollinearity in Regression	8
2.2 Multicollinearity and Outlier in Regression	
CHAPTER THREE: METHODOLOGY	11
3.1 Sources of Data	13
3.2 Method of Data Analysis	13
3.2.1 Ordinary Least Squares (OLS)	15
3.2.2 Jackknife Regression	16
3.2.3 Ridge Regression	17
3.2.4 Latent Root Regression	20
3.2.5 Detection of Multicollinearity	21
3.2.6 Detection of Outlier	22
3.3 Model Validation	22
3.3.1 Mean Square Error (MSE)	23
3.3.2 Coefficient of Determination () 2 R	